Epidemiology and outbreak investigation

By Pia D.M. MacDonald, PhD, MPH

This document provides a general summary of epidemiology and outbreak investigation.

Epidemiology is the study of the distribution and determinants of disease, risk factors and exposures, and health status in specified populations to improve population health. Epidemiology can study the health risk associated with specific exposures, it can be used as a tool to identify and assist in controlling epidemics, and it can monitor population rates of disease and exposures.

Epidemiology is an investigative process. When conducting epidemiologic studies, we are seeking to answer the questions who, what, when, where, why and how in order to understand the occurrence and causes of disease. As an example, using epidemiologic tools the relationship between smoking (exposure) and lung cancer (disease) was elucidated as was the association between absorbent tampon use (exposure) and toxic-shock syndrome (disease).

Descriptive epidemiology provides a systematic method for characterizing a health problem. By conducting descriptive epidemiology,

we can ensure that there is an understanding of the basic dimensions of a health problem. It also helps identify populations at higher risk for the health problem, provides information used for allocation of resources, and enables investigator to develop one or more hypotheses that can be tested.

Analytic epidemiology encompasses the formation of hypotheses, and designing studies to test those hypotheses. The hypothesis in an epidemiologic study explores the relationship between two main factors: an exposure and an outcome. Exposure refers to any factor that might influence one's risk of disease – this could be a virus, a pollutant, genetics, or many other things. Outcome refers to the disease in question, which is standardized by a case definition.

Within analytic epidemiology there are experimental and observational studies. Experimental studies can involve individuals or communities where an exposure could be assigned on an individual basis, as in clinical trials where new drug treatments are being tested, or the exposure could be assigned to an entire community, such as when health or educational campaigns are aimed at entire schools, neighborhoods, or cities. Observational studies do not include an intervention, they are based on observation alone. There are three main study designs that are used for observational epidemiology. These are the cross-sectional study, the cohort

study, and the case-control study. In all of these study designs, investigators gather data on an exposure and an outcome in a population of people.

Once an analytic study has been designed to test a hypothesis, and the data have been gathered, the data are analyzed to determine whether there is an association between the exposure and the outcome being studied. A measure of association is one product from an analytic study. The measure of association assesses the strength of a relationship between an exposure and the outcome of interest. To do this, the measure indicates how much more or less likely one group is to exhibit the outcome (for example a disease) compared to another group. One of the most commonly used measures of association is the relative risk (also known as the risk ratio). The other common measure of association is the odds ratio.

Another step in testing a hypothesis in an analytic study is to conduct a test of significance. A test of significance is an indication of the statistical confidence of the association that was observed. It answers the question "How likely is it that the observed association may be due to chance?" The 95% confidence interval and the p-value calculations assist with determining this. A 95% confidence interval is the range of values of the measure of association (the risk ratio or odds ratio) that has a 95% chance of containing the true relative risk or true odds ratio. P-values are a related way of

determining statistical significance. The p-value is a measure of how likely the observed association would be to occur by chance alone, if there was not really a true association.

An outbreak (or epidemic) is an increase in the observed number of cases of a disease or health problem compared with the expected number for a given place or among a specific group of people over a particular period of time. The number of cases it takes to indicate the occurrence of an outbreak varies according to the etiologic agent, the size and composition of the population, previous occurrence of the disease in the community (i.e., level of immunity due to previous disease or vaccination), and the time period of occurrence.

For rare diseases (e.g., smallpox, botulism, polio), even one case may be considered to be an outbreak. On the other hand, for common diseases (such as genital herpes or influenza), a very large number of cases may need to occur or cases may need to be uniquely related (e.g., occur in a well-circumscribed group or be of the same strain or genetic sequence) before they are considered to represent an outbreak.

Potential outbreaks come to the attention of public health officials in a variety of ways, including through health information systems, astute health

care providers, and affected citizens. Outbreaks may be detected through the routine and timely analysis of health information systems such as disease surveillance systems managed by state and local health departments and in selected health care settings.

Outbreak investigations include both descriptive and analytic epidemiology. They provide an opportunity to characterize a public health problem. For example, the investigation of 15 confirmed *Salmonella* cases in Minnesota uncovered a nationwide outbreak associated with consumption of Schwan's ice cream that may have gone undetected (1). An investigation may identify risk factors that are associated with infection that are preventable. Epidemiologic investigations of *Escherichia coli* O157:H7 outbreaks have identified consumption of foods such as pink hamburgers (2), unpasteurized apple juice (3), or alfalfa sprouts (4) that consumers may avoid to reduce their risk of illness. Outbreak investigations may provide new research insights into the disease even if no new cases are occurring. In 1986, identification of a small outbreak of chronic diarrheal illness of unknown etiology associated with consumption of unpasteurized milk has led to intense laboratory research into the presumed infectious agent of "Brainerd" diarrhea (5).

Once an outbreak is suspected, there are several steps that provide a systematic approach to investigation. The steps often occur simultaneously or may be repeated as new information is received. Epidemiology is used in an outbreak investigation to assist in determining who, what, when, where, why and how.

The outbreak investigation can be divided broadly into the preliminary phase, analytic study phase, and control and follow-up phase. In the preliminary phase, the background rate of disease is examined often through existing surveillance systems or medical record review to evaluate if there are indeed an unusual number of cases of disease in a specified population. A case definition(s) is established that includes a set of diagnostic criteria that must be fulfilled in order to identify a person as a case of a particular disease/syndrome. In some outbreaks the case definition is very straightforward, however, in others it can be quite complex. As a result, there can be laboratory-confirmed cases, clinically confirmed cases, probable cases, and possible/suspect cases. Case status can change throughout the investigation as more information (such as laboratory findings) becomes available about the cases.

To determine the scope of the outbreak (how many people have the same disease), active case finding, searching the population for patients with

the disease in question, is done by contacting, for example, public health departments, medical facilities, laboratories, and sometimes the general community.

Once potential cases are identified, the descriptive epidemiologic (person, place, time) information about them is collected, sometimes by interviewing the patients or their medical care providers, reviewing medical charts and laboratory findings, or examining routine surveillance data. The cases can be plotted on an epidemic curve. This epidemic curve is a histogram that provides a graphic depiction of the number of cases by date of illness. It is useful because it can provide information about the magnitude and type of outbreak (e.g. common source, point source or propagated), case outliers, time trend, and exposure and/or disease incubation period. Epidemiologic information about the cases such as date of illness, age, sex, race/ethnicity, residence, occupation, recent travel, or attendance at events is reviewed to see if there are any commonalities among them. If applicable and available, laboratory findings such as the genetic sequence or strain of the pathogen are important to examine as well. Using clinical (and laboratory if applicable) and epidemiological information, investigators can assess if case-patients are outbreak related or not.

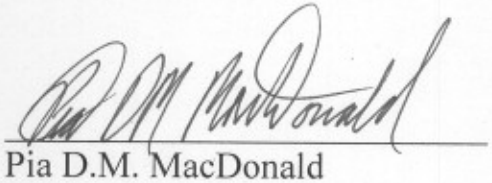
In the analytic study phase, the investigative team develops hypotheses regarding the cause of the outbreak. Knowledge of previous outbreaks and an understanding of the microbiology of the pathogen may provide useful clues, and open-ended interviews with several of the case patients may provide additional insight into the cause of the outbreak. Testing the list of hypotheses involves conducting analytic epidemiology studies such as cohort, case-control, or cross-sectional studies. During the analytic study, the case definition can be modified to increase specificity and reduce the chance of including cases of unrelated illness or no illness as outbreak-related cases.

In parallel to the epidemiologic investigation, environmental sampling should be considered. In a foodborne outbreak, for example, food and beverage samples should be collected as soon as possible. Ideally, microbiological testing of these samples should be focused by the results of the analytic study.

During the outbreak investigation, control measures should be taken to prevent further illness and prevent future outbreaks. This stage involves working with government regulators, industry, and health educators.

References

1. Hennessy TW, Hedberg CW, Slutsker L, et al. A national outbreak of *Salmonella enteritidis* infections from ice cream. The Investigation Team. *N Engl J Med* 1996;334:1281-6.
2. Bell BP, Goldoft M, Griffin PM, et al. A multistate outbreak of *Escherichia coli* O157:H7-associated bloody diarrhea and hemolytic uremic syndrome from hamburgers. The Washington experience. *JAMA* 1994;272:1349-53.
3. Cody SH, Glynn MK, Farrar JA, et al. An outbreak of *Escherichia coli* O157:H7 infection from unpasteurized commercial apple juice. *Ann Intern Med* 1999;130:202-9.
4. Breuer T, Benkel DH, Shapiro RL, et al. A multistate outbreak of *Escherichia coli* O157:H7 infections linked to alfalfa sprouts grown from contaminated seeds. *Emerg Infect Dis* 2001;7:977-82.
5. Osterholm MT, MacDonald KL, White KE, et al. An outbreak of a newly recognized chronic diarrhea syndrome associated with raw milk consumption. *JAMA* 1986;256:484-90.

Signature  Date 12/6/07

Pia D.M. MacDonald